GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: March 14, 2003, 03:18:01 ; Search time 54.1124 Seconds
(without alignments)
635.320 Million cell updates/sec

Title: US-09-698-781-3
Perfect score: 1436
Sequence: 1 MKQILHPALETTAMTLFPVL.....KHQLVRDSCKASCNCSNSIY 258

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1436	100.0	258	22	AAE02211	Human full-length
2	1363	94.9	245	22	AAM23992	Human EST encoded
3	1363	94.9	245	22	AAM24000	Human EST encoded
4	992	69.1	257	22	ABG06656	Novel human diagno
5	986	68.7	243	22	AAE13072	Homo sapiens (Hs)-
6	595	41.4	138	20	AAY44013	Human testis speci
7	572.5	39.9	168	22	ABG06655	Novel human diagno
8	507.5	35.3	137	20	AAY44012	Human testis speci
9	485.5	33.8	137	20	AAY44011	Rat sperm coating
10	354	24.7	71	20	AAY11989	Human 5' EST secre
11	344	24.0	255	23	AAE18962	Mouse testes-speci
12	322.5	22.5	219	21	AAB01400	Neuron-associated
13	318	22.1	302	21	AAB43408	Human cancer assoc
14	315	21.9	181	22	AAB64952	Gene 12 human secr
15	303	21.1	162	23	AAE21099	Human sperm-coatin
16	301.5	21.0	242	22	AAU12302	Human PRO7434 poly
17	300.5	20.9	242	21	AAB47998	Apple crisp protei
18	300.5	20.9	252	23	ABB80591	Human sbg1002620TI
19	300	20.9	91	22	ABG06658	Novel human diagno
20	295.5	20.6	241	22	AAM24028	Rat EST encoded pr
21	295.5	20.6	371	22	AAB93979	Human protein sequ
22	295.5	20.6	497	22	AAB74446	Human protease-inh
23	295.5	20.6	497	23	ABB80592	Human sbg1002620TI
24	295.5	20.6	497	23	AAE23977	Human LP095 secret
25	295.5	20.6	497	23	AAO21721	Human secreted pro
26	295.5	20.6	497	23	AAU83619	Human PRO protein,
27	295	20.5	463	22	AAB88408	Human membrane or
28	295	20.5	498	22	AAE10618	Human novel trypsi
29	294.5	20.5	477	22	AAE10617	Human novel trypsi
30	293.5	20.4	497	22	AAE10616	Human novel trypsi
31	291	20.3	266	22	AAM39716	Human polypeptide
32	291	20.3	347	22	AAM41502	Human polypeptide
33	291	20.3	463	20	AAY13392	Amino acid sequenc
34	291	20.3	463	21	AAB01373	Neuron-associated
35	291	20.3	463	21	AAY95343	Human PRO328 antit
36	291	20.3	463	22	AAU12351	Human PRO328 polyp
37	291	20.3	463	22	AAB80260	Human PRO328 prote
38	291	20.3	463	22	AAB53088	Human angiogenesis
39	290	20.2	198	16	AAR79914	Trypsin inhibitory
40	290	20.2	258	23	ABG61801	Prostate cancer-as
41	290	20.2	258	23	ABG61802	Prostate cancer-as
42	288.5	20.1	380	22	ABB71656	Drosophila melanog
43	288.5	20.1	431	21	AAB41816	Human ORFX ORF1580
44	288.5	20.1	500	20	AAY41738	Human PRO541 prote
45	288.5	20.1	500	21	AAB47999	Cocoa crisp protei

ALIGNMENTS

RESULT 1
AAE02211
ID AAE02211 standard; Protein; 258 AA.
XX
AC AAE02211;
XX
DT 31-JUL-2001 (first entry)
XX
DE Human full-length 36P1G3/SGP28 protein.
XX
KW Human; 36P1G3 clone; specific granule protein 28; SGP28; therapy;
KW prostate; colon; cancer; prognosis; vaccine; anticancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..32
FT /label= Signal_peptide
FT Protein 33..258
FT /note= "Human mature full-length 36P1G3/SGP28 protein"
FT Modified-site 26..31
FT /note= "N-myristoylation site; This region is
FT specifically referred in claim 18"
FT Modified-site 106..108
FT /note= "Protein kinase C phosphorylation site; This
FT region is specifically referred in claim 18"
FT Modified-site 108..114
FT /note= "Tyrosine kinase phosphorylation site; This
FT region is specifically referred in claim 18"
FT Modified-site 128..131
FT /note= "Casein kinase II phosphorylation site; This
FT region is specifically referred in claim 18"
FT Domain 148..158

```
FT                   /note= "Extracellular protein SCP signature sequence"
FT   Modified-site   164..169
FT                   /note= "N-myristoylation site; This region is
FT                   specifically referred in claim 18"
FT   Domain          179..190
FT                   /note= "Extracellular protein SCP signature sequence"
FT   Modified-site   188..193
FT                   /note= "N-myristoylation site; This region is
FT                   specifically referred in claim 18"
FT   Modified-site   201..206
FT                   /note= "N-myristoylation site; This region is
FT                   specifically referred in claim 18"
FT   Modified-site   206..209
FT                   /note= "Casein kinase II phosphorylation site; This
FT                   region is specifically referred in claim 18"
FT   Modified-site   214..219
FT                   /note= "N-myristoylation site; This region is
FT                   specifically referred in claim 18"
FT   Modified-site   231..233
FT                   /note= "Protein kinase C phosphorylation site; This
FT                   region is specifically referred in claim 18"
FT   Modified-site   252..255
FT                   /note= "N-glycosylation site; This
FT                   region is specifically referred in claim 18"
XX
PN   WO200131343-A2.
XX
PD   03-MAY-2001.
XX
PF   27-OCT-2000; 2000WO-US29607.
XX
PR   28-OCT-1999;   99US-0162610.
XX
PA   (UROG-) UROGENESYS INC.
XX
PI   Hubert RS,  Raitano AB,  Afar DEH,  Mitchell SC,  Faris M;
PI   Jakobovits A;
XX
DR   WPI; 2001-308685/32.
DR   N-PSDB; AAD06222.
XX
PT   Detecting cancers, particularly of prostate and colon, from
PT   overexpression of SGP28 protein, also methods for treating these
PT   cancers e.g. by vaccination with the protein -
XX
PS   Claim 16; Page 63; 102pp; English.
XX
CC   The present invention relates to methods and compositions for the
CC   diagnosis and therapy of prostate cancer which utilise human SGP28
CC   (specific granule protein 28) gene and proteins. The method involves
CC   detecting cancers, particularly of prostate and colon, from
CC   overexpression of SGP28 protein. The expression of SGP28, which is an
CC   extracellular protein is restricted to the prostate and ovary, and is
CC   markedly up-regulated in prostate tumours. SGP28 sequence is used for
CC   diagnosis (including in vivo imaging), staging, monitoring and prognosis
CC   of prostatic and colon cancer, and for assisting selection of therapy.
CC   Also SGP28-expressing cancers can be treated by administering a
CC   composition or vaccine that contains a vector expressing an antibody
CC   specific for SGP28 protein, nucleic acid encoding SGP28 protein or its
CC   fragments, polypeptides encoded by SGP28 gene and SGP28-specific antibody
CC   optionally conjugated to toxin or therapeutic agent. SGP28 gene product
CC   is also used as source of therapeutic antisense or ribozyme agents, as
CC   primers/probes for diagnosis or prognosis, to identify compounds that
CC   inhibit calcium entry into prostatic cells, for recombinant production
CC   of SGP28 peptides and for isolating related sequences. SGP28 protein and
CC   its fragments are used to raise specific antibodies (Ab) and to identify
CC   specific binding agents (potentially useful as therapeutic and
CC   diagnostic agents) and also potential anticancer agents. The present
CC   amino acid sequence is human full-length 36P1G3/SGP28 protein.
XX
SQ   Sequence   258 AA;

  Query Match             100.0%;  Score 1436;  DB 22;  Length 258;
  Best Local Similarity   100.0%;  Pred. No. 2e-126;
  Matches  258;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MKQILHPALETTAMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKQILHPALETTAMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHN 60

Qy      61 ELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPS 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPS 120

Qy     121 SWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYY 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYY 180

Qy     181 VCQYCPAGNWANRLYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKH 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 VCQYCPAGNWANRLYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKH 240

Qy     241 QLVRDSCKASCNCSNSIY 258
           ||||||||||||||||||
Db     241 QLVRDSCKASCNCSNSIY 258


RESULT 2
AAM23992
ID   AAM23992 standard; Protein; 245 AA.
XX
AC   AAM23992;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Human EST encoded protein SEQ ID NO: 1517.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   N-PSDB; AAH98651.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 20; Page 1047-1048; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a
CC   protein of the invention.
```

```
XX
SQ   Sequence   245 AA;

  Query Match            94.9%;  Score 1363;  DB 22;  Length 245;
  Best Local Similarity  99.6%;  Pred. No. 1.3e-119;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNM 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNM 60

Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      61 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSASSSWSQAIQSWFDEY 120

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 180

Qy     194 LYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCNC 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCNC 240

Qy     254 SNSIY 258
           |||||
Db     241 SNSIY 245


RESULT 3
AAM24000
ID   AAM24000 standard; Protein; 245 AA.
XX
AC   AAM24000;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Human EST encoded protein SEQ ID NO: 1525.
XX
KW   Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;
KW   tomato; monkey; dog; sea urchin; expressed sequence tag; EST;
KW   diagnostics; forensic test; gene mapping; genetic disorder;
KW   biodiversity; gene therapy; nutrition.
XX
OS   Homo sapiens.
XX
PN   WO200154477-A2.
XX
PD   02-AUG-2001.
XX
PF   25-JAN-2001; 2001WO-US02687.
XX
PR   25-JAN-2000; 2000US-0491404.
PR   17-JUL-2000; 2000US-0617746.
PR   03-AUG-2000; 2000US-0631451.
PR   15-SEP-2000; 2000US-0663870.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Zhou P,  Qian XB,  Wang Z,  Chen R,  Asundi V;
PI   Cao Y,  Drmanac RA,  Zhang J,  Werhman T;
XX
DR   WPI; 2001-476164/51.
DR   N-PSDB; AAH98659.
XX
PT   Isolated polypeptide for treatment of diseases, diagnostics, raising
PT   antibodies and research use -
XX
PS   Claim 20; Page 1051-1052; 1275pp; English.
XX
CC   The present invention provides the protein and coding sequences of novel
CC   proteins from a variety of organisms, including human, dog, cat, horse,
CC   cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea
CC   urchin and tomato. These were derived from expressed sequence tags (ESTs)
CC   from the organism of interest. They can be used in diagnostics,
CC   forensics, gene mapping, identification of mutations, to assess
CC   biodiversity and for nutritional purposes. The present sequence is a
CC   protein of the invention.
XX
SQ   Sequence   245 AA;

  Query Match            94.9%;  Score 1363;  DB 22;  Length 245;
  Best Local Similarity  99.6%;  Pred. No. 1.3e-119;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      14 MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNM 73
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNM 60

Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db      61 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSASSSWSQAIQSWFDEY 120

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 180

Qy     194 LYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCNC 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 LYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCNC 240

Qy     254 SNSIY 258
           |||||
Db     241 SNSIY 245


RESULT 4
ABG06656
ID   ABG06656 standard; Protein; 257 AA.
XX
AC   ABG06656;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #6647.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS70843.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 37015; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
```

```
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   257 AA;

  Query Match           69.1%;  Score 992;  DB 22;  Length 257;
  Best Local Similarity   72.0%;  Pred. No. 8.8e-85;
  Matches  177;  Conservative   25;  Mismatches   42;  Indels    2;  Gaps    2;

Qy      13 AMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARN 72
           || | || ||||  |||| || | |||||||||||| ||||||||||||||:||||||  |
Db      14 AMALLPV-LFLVTVLLPSLPA-EGKDPAFTALLTTQLQVQREIVNKHNELRKAVSPPASN 71

Qy      73 MLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDE 132
           ||||||::|   |||:|||:|  :||:|:|| || :|||||||| |:||| |||||:||
Db      72 MLKMEWSREVTTNAQRWANKCTLQHSDPEDRKTSTRCGENLYMSSDPTSWSSAIQSWYDE 131

Qy     133 YNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWAN 192
             || :|||||:||||||||||:||||:| |||| ||||||  ||||||||||||||  |
Db     132 ILDFVYGVGPKSPNAVVGHYTQLVWYSTYQVGCGIAYCPNQDSLKYYYVCQYCPAGNNMN 191

Qy     193 RLYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCN 252
           |   ||:|| ||| |||:|| ||||| |:|:|| ||| ||| |  |:|:|::: |||:|
Db     192 RKNTPYQQGTPCAGCPDDCDKGLCTNSCQYQDLLSNCDSLKNTAGCEHELLKEKCKATCL 251

Qy     253 CSNSIY 258
           | | ||
Db     252 CENKIY 257


RESULT 5
AAE13072
ID   AAE13072 standard; Protein; 243 AA.
XX
AC   AAE13072;
XX
DT   28-JAN-2002  (first entry)
XX
DE   Homo sapiens (Hs)-Tpx protein.
XX
KW   Angiogenesis; Ov-ASP; therapy; circulatory disorder; vascular disorder;
KW   congenital heart disease; myocardial disease; pericardial disease; wart;
KW   cerebrovascular ischaemia; veno-occlusive disease; myocardial ischaemia;
KW   coronary artery disease; diabetic retinopathy; inflammatory disease;
KW   wound healing; duodenal ulceration; rheumatoid arthritis; psoriasis;
KW   periodontitis; dermatological; cutaneous malignancy; Kaposi's sarcoma;
KW   pyogenic granuloma; cancer; onchocerciasis; River Blindness; neoplasia;
KW   vasotropic; cardiant; antiparasitic; opthalmological.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..22
FT                   /label= Signal_peptide
FT   Protein         23..243
FT                   /label= Mature_Hs-Tpx_protein
XX
PN   WO200174385-A1.
XX
PD   11-OCT-2001.
XX
PF   27-MAR-2001; 2001WO-US09798.
XX
PR   03-APR-2000; 2000US-0541759.
XX
PA   (NYBL-) NEW YORK BLOOD CENT INC.
PA   (UYCA-) UNIV CASE WESTERN RESERVE.
PA   (UABR-) UAB RES FOUND.
XX
PI   Lustigman S,  Pearlman E,  Unnasch TR;
XX
DR   WPI; 2001-662950/76.
XX
PT   Inducing angiogenesis in a tissue using the Ov-ASP protein isolated
PT   from the nematode Onchocera volvulus is useful to treat circulatory or
PT   vascular disease such as ischemia  -
XX
PS   Disclosure; Fig 1; 37pp; English.
XX
CC   The present invention relates to a method for inducing angiogenesis in a
CC   tissue. The method comprising contacting the tissue with Ov-ASP. The Ov-
CC   ASP molecules are used to treat circulatory or vascular disorders,
CC   particularly ischaemia, congenital heart disease, myocardial disease or
CC   pericardial disease, more particularly cerebrovascular ischaemia, veno-
CC   occlusive disease or myocardial ischaemia, especially coronary artery
CC   disease. The invention is also used to treat cancer, diabetic
CC   retinopathy and inflammatory disease. Angiogenesis is also central to a
CC   number of pathological processes, including abnormalities of wound
CC   healing in diseases such as diabetes and duodenal ulceration; chronic
CC   inflammatory disorders such as rheumatoid arthritis, psoriasis and
CC   periodontitis; dermatological conditions such as cutaneous malignancy,
CC   Kaposi's sarcoma, pyogenic granulomas and warts. Anti-Ov-ASP factors are
CC   useful to treat onchocerciasis (River Blindness) or benign or malignant
CC   neoplasia. The present sequence is homo sapiens (Hs)-Tpx protein.
XX
SQ   Sequence   243 AA;

  Query Match           68.7%;  Score 986;  DB 22;  Length 243;
  Best Local Similarity   71.4%;  Pred. No. 3e-84;
  Matches  175;  Conservative   26;  Mismatches   42;  Indels    2;  Gaps    2;

Qy      14 MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNM 73
           | | || ||||  |||| || | |||||||||||| ||||||||||:|:|||||| ||
Db       1 MALLPV-LFLVTVLLPSLPA-EGKDPAFTALLTTQLQVQREIVNKHNEIRKAVSPPASNM 58

Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |||||::|   |||:|||:|  :||:|:|| || :|||||||| |:||| |||||:||
Db      59 LKMEWSREVTTNAQRWANKCTLQHSDPEDRKTSTRCGENLYMSSDPTSWSSAIQSWYDEI 118

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
            || :|||||:||||||||||:||||:| |||| ||||||  ||||||||||||||  ||
Db     119 LDFVYGVGPKSPNAVVGHYTQLVWYSTYQVGCGIAYCPNQDSLKYYYVCQYCPAGNNMNR 178

Qy     194 LYVPYEQGAPCASCPDNCDDGLCTNGCKYEDLYSNCKSLKLTLTCKHQLVRDSCKASCNC 253
              ||:|| ||| |||:|| ||||| |:|:|| ||| ||| |  |:|:|::: |||:| |
Db     179 KNTPYQQGTPCAGCPDDCDKGLCTNSCQYQDLLSNCDSLKNTAGCEHELLKEKCKATCLC 238

Qy     254 SNSIY 258
             | ||
Db     239 ENKIY 243


RESULT 6
AAY44013
ID   AAY44013 standard; Protein; 138 AA.
XX
AC   AAY44013;
XX
DT   21-DEC-1999  (first entry)
XX
```

```
DE   Human testis specific protein #2.
XX
KW   Prediction; secondary structure; alignment; evolutionary conservation;
KW   homology; periodicity; co-variation analysis; antigenic site;
KW   site directed mutagenesis; interaction.
XX
OS   Homo sapiens.
XX
PN   US5958784-A.
XX
PD   28-SEP-1999.
XX
PF   25-MAR-1992;   92US-0857224.
XX
PR   25-MAR-1992;   92US-0857224.
XX
PA   (BENN/) BENNER S A.
XX
PI   Benner SA;
XX
DR   WPI; 1999-570766/48.
XX
PT   Predicting the folded structure of proteins -
XX
PS   Disclosure; Column 389-390; 113pp; English.
XX
CC   Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC   predicting the folded structure of proteins, by aligning sequences of
CC   homologous proteins and using patterns of evolutionarily conserved and
CC   varied sequences to assign positions. Positions in the alignment are
CC   assigned to the surface or inside of the folded structure, active sites,
CC   and parsing segments. Secondary structural units are assigned by
CC   identifying periodicity in the assignments, and assembled into globular
CC   form using distance constraints imposed by disulfide bridges, active
CC   site assignments and co-variation analysis. The predicted secondary
CC   structures are useful for identifying antigenic sites on a protein
CC   molecule, as guides for site directed mutagenesis studies, and for
CC   understanding the interaction of a protein with other molecules.
XX
SQ   Sequence   138 AA;

  Query Match             41.4%;  Score 595;  DB 20;  Length 138;
  Best Local Similarity   71.7%;  Pred. No. 6.7e-48;
  Matches   99;  Conservative   16;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |||||::|   |||:|||:|  :||:|:|| || :||||||||| |:||| |||||:||
Db       1 LKMEWSREVTTNAQRWANKCTLQHSDPEDRKTSTRCGENLYMSSDPTSWSSAIQSWYDEI 60

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
           || :|||||:|||||||||:||||:| |||| ||||||  ||||||||||||||  ||
Db      61 LDFVYGVGPKSPNAVVGHYTQLVWYSTYQVGCGIAYCPNQDSLKYYYVCQYCPAGNNMNR 120

Qy     194 LYVPYEQGAPCASCPDNC 211
             ||:|| ||| |||:|
Db     121 KNTPYQQGTPCAGCPDDC 138


RESULT 7
ABG06655
ID   ABG06655 standard; Protein; 168 AA.
XX
AC   ABG06655;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Novel human diagnostic protein #6646.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   N-PSDB; AAS70842.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 20; SEQ ID No 37014; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. ABG00010-ABG30377 represent novel human
CC   diagnostic amino acid sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence   168 AA;

  Query Match             39.9%;  Score 572.5;  DB 22;  Length 168;
  Best Local Similarity   58.9%;  Pred. No. 1.1e-45;
  Matches  113;  Conservative   14;  Mismatches   26;  Indels   39;  Gaps    3;

Qy      13 AMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARN 72
           || | || |||| |||| || | |||||||||||| ||||||||||||||:|||||| |
Db      14 AMALLPV-LFLVTVLLPSLPA-EGKDPAFTALLTTQLQVQREIVNKHNELRKAVSPPASN 71

Qy      73 MLKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDE 132
           ||||||::|   |||:|||:|  :||:|:|| || :||||||||| |:||| |||||:||
Db      72 MLKMEWSREVTTNAQRWANKCTLQHSDPEDRKTSTRCGENLYMSSDPTSWSSAIQSWYDE 131

Qy     133 YNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWAN 192
           || :|||||:|                                    |||||||  |
Db     132 ILDFVYGVGPKSP-----------------------------------QYCPAGNNMN 154

Qy     193 RLYVPYEQGAPC 204
           |   ||:|| ||
Db     155 RKNTPYQQGTPC 166


RESULT 8
AAY44012
ID   AAY44012 standard; Protein; 137 AA.
XX
AC   AAY44012;
```

XX
DT 21-DEC-1999 (first entry)
XX
DE Human testis specific protein #1.
XX
KW Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.
XX
OS Homo sapiens.
XX
PN US5958784-A.
XX
PD 28-SEP-1999.
XX
PF 25-MAR-1992; 92US-0857224.
XX
PR 25-MAR-1992; 92US-0857224.
XX
PA (BENN/) BENNER S A.
XX
PI Benner SA;
XX
DR WPI; 1999-570766/48.
XX
PT Predicting the folded structure of proteins
XX
PS Disclosure; Column 387-388; 113pp; English.
XX
CC Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and parsing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active
CC site assignments and co-variation analysis. The predicted secondary
CC structures are useful for identifying antigenic sites on a protein
CC molecule, as guides for site directed mutagenesis studies, and for
CC understanding the interaction of a protein with other molecules.
XX
SQ Sequence 137 AA;

Query Match 35.3%; Score 507.5; DB 20; Length 137;
Best Local Similarity 61.6%; Pred. No. 1.1e-39;
Matches 85; Conservative 20; Mismatches 32; Indels 1; Gaps 1;

```
Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |||||: :|  |||||||:|   ||:  ||  :::|||||||||: |: ||  ||||::|
Db       1 LKMEWSIQATTNAQKWANKCILEHSSKDDRKINIRCGENLYMSTDPTLWSTVIQSWYNEN 60

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
             || :||| | ||: ||||||:|||||: :||| ||||||  |||:||| ||| ||    :
Db      61 EDFVYGVGAK-PNSAVGHYTQLVWYSSFKIGCGIAYCPNQDNLKYFYVCHYCPMGNNVMK 119

Qy     194 LYVPYEQGAPCASCPDNC 211
              ||:|| ||||||:||
Db     120 KSTPYQQGTPCASCPNNC 137
```

RESULT 9
AAY44011
ID AAY44011 standard; Protein; 137 AA.
XX
AC AAY44011;
XX
DT 21-DEC-1999 (first entry)
XX
DE Rat sperm coating glycoprotein.
XX
KW Prediction; secondary structure; alignment; evolutionary conservation;
KW homology; periodicity; co-variation analysis; antigenic site;
KW site directed mutagenesis; interaction.
XX
OS Bos taurus.
XX
PN US5958784-A.
XX
PD 28-SEP-1999.
XX
PF 25-MAR-1992; 92US-0857224.
XX
PR 25-MAR-1992; 92US-0857224.
XX
PA (BENN/) BENNER S A.
XX
PI Benner SA;
XX
DR WPI; 1999-570766/48.
XX
PT Predicting the folded structure of proteins
XX
PS Disclosure; Column 385-388; 113pp; English.
XX
CC Sequences AAY43902-Y44015 represent proteins used in a novel method of
CC predicting the folded structure of proteins, by aligning sequences of
CC homologous proteins and using patterns of evolutionarily conserved and
CC varied sequences to assign positions. Positions in the alignment are
CC assigned to the surface or inside of the folded structure, active sites,
CC and parsing segments. Secondary structural units are assigned by
CC identifying periodicity in the assignments, and assembled into globular
CC form using distance constraints imposed by disulfide bridges, active
CC site assignments and co-variation analysis. The predicted secondary
CC structures are useful for identifying antigenic sites on a protein
CC molecule, as guides for site directed mutagenesis studies, and for
CC understanding the interaction of a protein with other molecules.
XX
SQ Sequence 137 AA;

Query Match 33.8%; Score 485.5; DB 20; Length 137;
Best Local Similarity 60.1%; Pred. No. 1.2e-37;
Matches 83; Conservative 18; Mismatches 36; Indels 1; Gaps 1;

```
Qy      74 LKMEWNKEAAANAQKWANQCNYRHSNPKDRMTSLKCGENLYMSSAPSSWSQAIQSWFDEY 133
           |::||: :|  |||||||:| | ||  : | |:|||||||:|:: |:|||  || |:||
Db       1 LRVEWDHDAYVNAQKWANRCIYNHSPLQHRTTTLKCGENLFMANYPASWSSVIQDWYDES 60

Qy     134 NDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAGNWANR 193
             || || |||     ||||||||| |::|| || | ||:|  |||:||| ||| ||:  |
Db      61 LDFVFGFGPKKVGVKVGHYTQVVWNSTFLVACGVAECPDQP-LKYFYVCHYCPGGNYVGR 119

Qy     194 LYVPYEQGAPCASCPDNC 211
           || || :| || ||| ||
Db     120 LYSPYTEGEPCDSCPGNC 137
```

RESULT 10
AAY11989
ID AAY11989 standard; Protein; 71 AA.
XX
AC AAY11989;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human 5' EST secreted protein SEQ ID No: 589.
XX
KW Human; secreted protein; EST; expressed sequence tag; diagnosis;
KW forensic; gene therapy; chromosome mapping; signal peptide; prostate;
KW upstream regulatory sequence; cytokine activity; cell proliferation;
KW differentiation; haematopoiesis regulation; tissue growth regulation;
KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
KW thrombolytic; anti-inflammatory; tumour inhibition.
XX
OS Homo sapiens.
XX

```
PN   WO9906550-A2.
XX
PD   11-FEB-1999.
XX
PF   31-JUL-1998;   98WO-IB01232.
XX
PR   01-AUG-1997;   97US-0905144.
XX
PA   (GEST ) GENSET.
XX
PI   Duclert A,  Dumas Milne Edwards J,  Lacroix B;
XX
DR   WPI; 1999-153780/13.
DR   N-PSDB; AAX40711.
XX
PT   New isolated prostate-derived nucleic acids - used to develop
PT   products which may have cytokine, immune regulatory, haematopoiesis
PT   regulating, anti-inflammatory or tumour inhibition activity
XX
PS   Claim 34; Page 672; 675pp; English.
XX
CC   AAX40438 to AAX40715 represent 5' expressed sequence tags (ESTs) for
CC   human secreted proteins expressed in prostate, and encode the proteins
CC   given in AAY11716 to AAY11993 respectively. The proteins given represent
CC   the signal peptide and an N-terminal fragment of a secreted protein. The
CC   nucleic acid sequences can be used for producing secreted human gene
CC   products. They can also be used to develop products for diagnosis and
CC   therapy. The proteins obtained may have cytokine activity, cell
CC   proliferation and differentiation activity, haematopoiesis regulating
CC   activity, tissue growth regulating activity, reproductive hormone
CC   regulating activity, chemotactic/chemokinetic activity, haemostatic and
CC   thrombolytic activity, receptor/ligand activity, anti-inflammatory
CC   activity, tumour inhibition activity or other activities. The products
CC   can be used in forensic, gene therapy and chromosome mapping procedures.
CC   The sequences can also be used for obtaining corresponding promoter
CC   sequences. The nucleic acids encoding the signal peptides can be used for
CC   directing extracellular secretion of a polypeptide or the insertion of a
CC   polypeptide into a membrane, or importing a polypeptide into a cell.
XX
SQ   Sequence   71 AA;

  Query Match             24.7%;  Score 354;  DB 20;  Length 71;
  Best Local Similarity   98.6%;  Pred. No. 1.1e-25;
  Matches   70;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MKQILHPALETTAMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHN 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MKQILHPALETTAMTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHN 60

Qy      61 ELRRAVSPPAR 71
           ||||||||||:
Db      61 ELRRAVSPPAK 71


RESULT 11
AAE18962
ID   AAE18962 standard; Protein; 255 AA.
XX
AC   AAE18962;
XX
DT   21-MAY-2002  (first entry)
XX
DE   Mouse testes-specific, vespid and pathogenic protein (RTVP).
XX
KW   Mouse; testes-specific, vespid and pathogenic protein; RTVP; therapy;
KW   anti-neoplastic; prostatic neoplasia; prostate carcinoma; cytokine;
KW   metastatic disease; neoplastic disease; immune system; growth factor;
KW   cytostatic.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..16
FT                   /label= Signal_peptide
FT   Protein         17..255
FT                   /note= "Mouse mature RTVP protein"
FT   Modified-site   90..92
FT                   /note= "N-glycosylation site"
FT   Domain          135..144
FT                   /note= "Extracellular protein signature motif 1"
FT   Domain          160..170
FT                   /note= "Extracellular protein signature motif 2"
FT   Misc-difference 195
FT                   /note= "Encoded by TG"
FT   Misc-difference 210
FT                   /note= "Encoded by GCAT"
FT   Domain          222..244
FT                   /note= "Transmembrane domain"
XX
PN   WO200206344-A2.
XX
PD   24-JAN-2002.
XX
PF   08-JUN-2001; 2001WO-US18487.
XX
PR   08-JUN-2000; 2000US-209989P.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI   Thompson TC,  Ren C;
XX
DR   WPI; 2002-195804/25.
DR   N-PSDB; AAD30356.
XX
PT   Novel testes-specific, vespid and pathogenic polypeptide useful for
PT   treating and preventing prostatic neoplastic diseases, such as
PT   prostatic carcinoma and metastatic carcinoma, has antineoplastic
PT   activity -
XX
PS   Claim 22; Fig 1B; 72pp; English.
XX
CC   The invention relates to a gene encoding non-human testes-specific,
CC   vespid and pathogenic protein (RTVP) having anti-neoplastic activity.
CC   The invention further relates to compositions and methods based on RTVP
CC   for the treatment, prevention and detection of prostatic neoplasia such
CC   as prostate carcinoma and associated metastatic disease. Diagnostic kit
CC   comprising RTVP protein is useful for the detection of neoplastic
CC   disease. Composition comprising RTVP protein is useful in the diagnosis,
CC   studying and treatment of prostatic neoplasia such as prostatic carcinoma
CC   and associated metastatic disease. It is also useful for stimulating
CC   immune system e.g. cytokines and growth factors in a patient. The present
CC   sequence is mouse RTVP protein.
XX
SQ   Sequence   255 AA;

  Query Match             24.0%;  Score 344;  DB 23;  Length 255;
  Best Local Similarity   33.6%;  Pred. No. 5.3e-24;
  Matches   88;  Conservative   37;  Mismatches   81;  Indels   56;  Gaps   12;

Qy      19 VLLFLVAGLLPSFPANEDKDPAFTALL---TTQTQVQREIVNKHNELRRAVSPPARNMLK 75
           |:| ::  :  |  ::     :|||      |     :| |  ||:||  |||||||||
Db       3 VILAVIVWMASSVSSS-----SFTASTLPDITNEDFIKECVQVHNQLRSKVSPPARNMLY 57

Qy      76 MEWNKEAAANAQKWANQCNYRHSNPK------DRMTSLKCGENLYMSS-APSSWSQAIQS 128
           | |: : |  |: |    | ::| ||:          |:|  |||::: | :  | | || :
Db      58 MSWDPKLAQIAKAWTKSCEFKH-NPQLHSRIHPNFTAL--GENIWLGSLSIFSVSSAISA 114

Qy     129 WFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKYYYVCQYCPAG 188
           |:::|   :||    :    | |||||||  || :||    |||       ::| | |||
Db     115 WYEEIKHYDFST--RKCRHVCGHYTQVVWADSYKLGCAVQLCPNGA----NFICDYGPAG 168

Qy     189 NWANRLYVPYEQGAPCASCP--DNCDDGLCTN--------------------GCKYEDL 225
           |:      ||:||| |: ||  | | : || |                     :|  |
Db     169 NYPT---WPYKQGATCSDCPKDDKCLNSLCINPRRDQVSRYYSVDYPDWPIYLRNRYTSL 225

Qy     226 YSNCKSLKL------TLTCKHQ 241
```

```
          :   ||: |      |:  ||:
Db    226 FLIAKSVLLLLSVIITIWVKHK 247


RESULT 12
AAB01400
ID   AAB01400 standard; Protein; 219 AA.
XX
AC   AAB01400;
XX
DT   20-OCT-2000  (first entry)
XX
DE   Neuron-associated protein.
XX
KW   Neuron associated protein; NEUAP; neurological disorder; epilepsy;
KW   ischemic cerebrovascular disease; stroke; cerebral neoplasm;
KW   Alzheimer's disease; Pick's disease; Huntington's disease;
KW   dementia; Parkinson's disease; demyelinating disease; meningitis;
KW   prion disease; kuru; Creutzfeldt-Jakob disease; neurofibromatosis;
KW   cerebral palsy; muscular dystrophy; central nervous system; CNS;
KW   peripheral nervous system; PNS; myopathy; schizophrenia;
KW   actinic keratosis; arteriosclerosis; atherosclerosis; bursitis;
KW   cirrhosis; hepatitis; mixed connective tissue disease; MCTD;
KW   myelofibrosis; paroxysmal nocturnal haemoglobinuria; cancer;
KW   autoimmune disease; inflammation; acquired immunodeficiency syndrome;
KW   AIDS; Addison's disease; adult respiratory distress syndrome;
KW   allergy; ankylosing spondylitis; amyloidosis; anaemia; asthma;
KW   Werner syndrome, trauma; human.
XX
OS   Homo sapiens.
XX
PN   WO200034477-A2.
XX
PD   15-JUN-2000.
XX
PF   10-DEC-1999;   99WO-US30408.
XX
PR   11-DEC-1998;   98US-0210083.
PR   11-DEC-1998;   98US-9123456.
PR   09-FEB-1999;   99US-0119365.
PR   16-MAR-1999;   99US-0124687.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Tang YT,  Yue H,  Baughn MR,  Hillman JL,  Lal P,  Au-young J,  Yang J;
PI   Lu DAM,  Azimzai Y;
XX
DR   WPI; 2000-423423/36.
XX
PT   New human neuron-associated proteins and polynucleotides encoding them,
PT   useful for diagnosis, treatment and prevention of cell proliferative
PT   disorders including cancer, neuronal and neurological disorders
XX
PS   Disclosure; Page 144-145; 145pp; English.
XX
CC   Human neuron-associated proteins (NEUAP) can be used for for
CC   treating or preventing a disorder associated with decreased
CC   expression or activity of NEUAP. Antagonists of NEUAP are useful for
CC   treating or preventing disorder associated with increased expression
CC   or activity of NEUAP. NEUAP or their fragments or derivatives are
CC   useful for treating neurological disorder such as epilepsy, ischemic
CC   cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's
CC   disease, Pick's disease, Huntington's disease, dementia and
CC   Parkinson's disease. NEUAPs are also useful for treating other
CC   demyelinating diseases, bacterial and viral meningitis, prion
CC   diseases including kuru, Creutzfeldt-Jakob disease, nutritional and
CC   metabolic diseases of the nervous system, neurofibromatosis, other
CC   developmental disorders of the central nervous system, cerebral
CC   palsy, neuroskeletal disorders, autonomic nervous system disorders,
CC   cranial nerve disorders, spinal cord diseases, muscular dystrophy and
CC   other neuromuscular disorders, peripheral nervous system disorders,
CC   inherited, metabolic, endocrine, and toxic myopathies, mental
CC   disorders including mood, anxiety and schizophrenic disorders, a cell
CC   proliferative disorder such as actinic keratosis, arteriosclerosis,
CC   atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective
CC   tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal
CC   haemoglobinuria, cancers of the adrenal gland, bladder, bone,
CC   bone marrow, brain, breast, cervix, and an autoimmune/inflammatory
CC   disorder such as acquired immunodeficiency syndrome (AIDS), Addison's
CC   disease, adult respiratory distress syndrome, allergies, ankylosing
CC   spondylitis, amyloidosis, anemia, asthma, Werner syndrome,
CC   complications of cancer, hemodialysis, and extracorporeal circulation,
CC   viral, bacterial, fungal parasitic, protozoal, and helminthic
CC   infections, and trauma. This protein was designated g847722.
XX
SQ   Sequence   219 AA;

  Query Match             22.5%;  Score 322.5;  DB 21;  Length 219;
  Best Local Similarity   36.5%;  Pred. No. 4.5e-22;
  Matches   77;  Conservative   30;  Mismatches   63;  Indels   41;  Gaps   11;

Qy      25 AGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHNELRRAVSPPARNMLKMEWNKEAAA 84
           | :||    |||               :: |  ||: |  | | | :|| | |:   |
Db      12 ANILPDI-ENED-------------FIKDCVRIHNKFRSEVKPTASDMLYMTWDPALAQ 56

Qy      85 NAQKWANQCNYRHS---NPKDRM----TSLKCGENLYMSSAP-SSWSQAIQSWFDEYNDF 136
             |: ||: | : |:    |  ::    |||  |||::  | |  | | || :|:||  |:
Db      57 IAKAWASNCQFSHNTRLKPPHKLHPNFTSL--GENIWTGSVPIFSVSSAITNWYDEIQDY 114

Qy     137 DFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPNQKVLKY-------YYVCQYCPAGN 189
           :|    :    | ||||||||  || |||   :||  ||  :       :::| | | ||
Db     115 NFKT--RICKKVCGHYTQVVWADSYKVGCAVQFCP--KVSGFDALSNGAHFICNYGPGGN 170

Qy     190 WANRLYVPYEQGAPCASCPDN--CDDGLCTN 218
           :      ||::|| |::||:|  | | || |
Db     171 YPT---WPYKRGATCSACPNNDKCLDNLCVN 198


RESULT 13
AAB43408
ID   AAB43408 standard; Protein; 302 AA.
XX
AC   AAB43408;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated protein sequence SEQ ID NO:853.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   N-PSDB; AAC77617.
XX
```

```
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 11; Page 1406-1407; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence   302 AA;

  Query Match             22.1%;  Score 318;  DB 21;  Length 302;
  Best Local Similarity   35.3%;  Pred. No. 1.8e-21;
  Matches   83;  Conservative   31;  Mismatches   79;  Indels   42;  Gaps   12;

Qy       2 KQILHPALETTA-MTLFPVLLFLVAGLLPSFPANEDKDPAFTALLTTQTQVQREIVNKHN 60
Db      34 RQSMRVTLATIAWMVSFVSNYSHTANILPDI-ENED--------------FIKDCVRIHN 78

Qy      61 ELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHS---NPKDRM----TSLKCGENL 113
Db      79 KFRSEVKPTASDMLYMTWDPALAQIAKAWASNCQFSHNTRLKPPHKLHPNFTSL--GENI 136

Qy     114 YMSSAP-SSWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVGCGNAYCPN 172
Db     137 WTGSVPIFSVSSAITNWYDEIQDYDFKT--RICKKVCGHYTQVVWADSYKVGCAVQFCP- 193

Qy     173 QKVLKY-------YYVCQYCPAGNWANRLYVPYEQGAPCASCPDN--CDDGLCTN 218
Db     194 -KVSGFDALSNGAHFICNYGPGGNYPT---WPYKRGATXSACPNNDKCLDNLCVN 244


RESULT 14
AAB64952
ID   AAB64952 standard; Protein; 181 AA.
XX
AC   AAB64952;
XX
DT   23-MAR-2001  (first entry)
XX
DE   Gene 12 human secreted protein homologous amino acid sequence #130.
XX
KW   Human; secreted protein; diagnosis; immunomodulatory; antisclerotic;
KW   dermatological; immunosuppressive; antiinflammatory; anti-HIV;
KW   immunostimulant; cytostatic; cardiant; vascular; anti-angiogenic;
KW   ophthalmological; neuroprotectant; nootropic; anticonvulsant; vulnerary;
KW   antialzheimers; antiparkinsonian; antimicrobial; immune disorder;
KW   multiple sclerosis; systemic lupus erythematosus; HIV; infection;
KW   hyperproliferative disorder; cancer; Gaucher's disease; wound healing;
KW   cardiovascular disease; Scimitar syndrome; Chaga's cardiomyopathy;
KW   coronary arteriosclerosis; angiogenic disorder; diabetic retinopathy;
KW   corneal graft neovascularisation; neurological disorder; regeneration;
KW   Huntington's chorea; Alzheimer's disease; Parkinson's disease;
KW   infectious disease; chemotaxis.
XX
OS   Homo sapiens.
XX
PN   WO200076530-A1.
XX
PD   21-DEC-2000.
XX
PF   01-JUN-2000; 2000WO-US14933.
XX
PR   11-JUN-1999;   99US-0138572.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (ROSE/) ROSEN C A.
XX
PI   Rosen CA,  Ruben SM,  Komatsoulis GA;
XX
DR   WPI; 2001-071147/08.
XX
PT   Nucleic acids encoding 49 human secreted polypeptides, useful for
PT   preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT   disease and diabetic retinopathy -
XX
PS   Disclosure; Page 528-529; 554pp; English.
XX
CC   The polynucleotide sequences given in AAF33213 to AAF33261 encode the
CC   human secreted proteins given in AAB64882 to AAB64930. AAB64931 to
CC   AAB64991 represent human secreted polypeptide sequences and proteins
CC   homologous to them, which are given in the exemplification of the present
CC   invention. Human secreted proteins have activities based on the tissues
CC   and cells the genes are expressed in. Examples of activities include:
CC   immunomodulatory; antisclerotic; dermatological; immunosuppressive;
CC   antiinflammatory; anti-HIV; immunostimulant; cytostatic; cardiant;
CC   vascular; antimicrobial; anti-angiogenic; ophthalmological;
CC   neuroprotectant; anticonvulsant; nootropic; antialzheimers;
CC   antiparkinsonian; and vulnerary. The polynucleotides and polypeptides can
CC   be used in the prevention, diagnosis and treatment of diseases associated
CC   with inappropriate polypeptide expression. Disorders that may be
CC   prevented, diagnosed and/or treated by the above methods include immune
CC   disorders (e.g. multiple sclerosis, systemic lupus erythematosus and
CC   human immuno-deficiency virus (HIV) infections), hyperproliferative
CC   disorders (e.g. cancers and Gaucher's disease), cardiovascular diseases
CC   (e.g. Scimitar syndrome, Chaga's cardiomyopathy and coronary
CC   arteriosclerosis), angiogenic disorders (e.g. corneal graft
CC   neovascularisation and diabetic retinopathy), neurological disorders
CC   (e.g. Huntington's chorea, Alzheimer's disease and Parkinson's disease),
CC   infectious diseases and/or for promoting wound healing, regeneration and
CC   /or chemotaxis. AAF33204 to AAF33212 and AAB64881 represent sequences
CC   used in the exemplification of the present invention.
XX
SQ   Sequence   181 AA;

  Query Match             21.9%;  Score 315;  DB 22;  Length 181;
  Best Local Similarity   38.7%;  Pred. No. 1.8e-21;
  Matches   70;  Conservative   29;  Mismatches   56;  Indels   26;  Gaps    9;

Qy      53 REIVNKHNELRRAVSPPARNMLKMEWNKEAAANAQKWANQCNYRHS---NPKDRM----T 105
Db      10 KDCVRIHNKFRSEVKPTASDMLYMTWDPALAQIAKAWASNCQFSHNTRLKPPHKLHPNFT 69

Qy     106 SLKCGENLYMSSAP-SSWSQAIQSWFDEYNDFDFGVGPKTPNAVVGHYTQVVWYSSYLVG 164
Db      70 SL--GENIWTGSVPIFSVSSAITNWYDEIQDYNFKT--RICKKVCGHYTQVVWADSYKVG 125

Qy     165 CGNAYCPNQKVLKY-------YYVCQYCPAGNWANRLYVPYEQGAPCASCPDN--CDDGL 215
Db     126 CAVQFCP--KVSGFDALSNGAHFICNYGPGGNYPT---WPYKRGATCSACPNNDKCLDNL 180

Qy     216 C 216
Db     181 C 181


RESULT 15
AAE21099
```

ID AAE21099 standard; Protein; 162 AA.
XX
AC AAE21099;
XX
DT 01-JUL-2002 (first entry)
XX
DE Human sperm-coating protein (SCP)-like domain consensus sequence.
XX
KW Human; haematopoiesis; clotting; kidney failure; wound healing; cancer;
KW neoplasia; pancreatic disorder; pancreatitis; cerebrovascular disease;
KW heart disorder; ischaemic heart disease; neuroprotective; vulnerary;
KW cardiovascular disorder; ischaemic heart disease; immunosuppressive;
KW glomerular disease; glomerulonephritis; uterine disorder; hyperplasia;
KW fetal spleen; prostate disorder; inflammatory disease; Crohn's disease;
KW proliferative disorder; gynaecological; haemostatic; antibacterial;
KW systemic lupus erythematosus; immunodeficiency disorder; antiasthmatic;
KW cytostatic; nephrotropic; antidiabetic; cerebroprotective; tranquilliser;
KW hypotensive; tumour; injury; trauma; antianginal; vasotropic; antiulcer;
KW apoptotic disorder; rheumatoid arthritis; cardiant; renal disorder;
KW hepatotropic; antipsoriatic; antiallergic; dermatological; virucide;
KW sperm-coating protein; SCP.
XX
OS Homo sapiens.
XX
PN US2002028508-A1.
XX
PD 07-MAR-2002.
XX
PF 21-FEB-2001; 2001US-0790264.
XX
PR 23-APR-1998; 98US-0065363.
PR 23-APR-1998; 98US-0065661.
PR 22-JUN-1998; 98US-0102705.
PR 29-JUL-1998; 98US-0124538.
PR 23-APR-1999; 99US-0298531.
PR 22-JUN-1999; 99US-0337930.
PR 29-JUL-1999; 99US-0363630.
XX
PA (HOLT/) HOLTZMAN D A.
PA (GOOD/) GOODEARL A D J.
PA (MCCA/) MCCARTHY S A.
XX
PI Holtzman DA, Goodearl ADJ, Mccarthy SA;
XX
DR WPI; 2002-303420/34.
XX
PT Novel TANGO polypeptides and nucleic acid molecules useful as
PT modulating agents in regulating cellular processes and for diagnosing
PT and treating heart, liver, lung, kidney, inflammatory and cellular
PT proliferative disorders -
XX
PS Example 3; Fig 2A; 138pp; English.
XX
CC The invention relates to nucleic acids encoding a variety of proteins
CC human T139 (TANGO-139), T125 (TANGO-125), T110 (TANGO-110), murine T175
CC (TANGO-175), human T175 or murine WDNM-2, having diagnostic, preventive,
CC therapeutic and other uses. Polypeptide of the invention has the ability
CC to inhibit a proteinase activity, to modulate cell-cell interactions,
CC haematopoiesis and the ability to modulate clotting. Polypeptide and
CC polynucleotide of the invention are useful for diagnosing and treating
CC disorder characterised by their aberrant expression or activity. The
CC antibodies are useful as modulating agents in regulating a variety of
CC cellular processes e.g. cell proliferation and/or cell differentiation.
CC TANGO-139 is useful for treating kidney defects such as kidney failure,
CC TANGO-125 is useful in wound healing and for treating cancer, TANGO-110
CC is useful for treating neoplasia, TANGO-177 or WDNM-2 is useful for
CC treating cancer, are useful to treat pancreatic disorders, such as
CC pancreatitis, cerebrovascular disease, and tumours, and injury or trauma
CC to the brain. TANG0-125, 110, 175 molecules treat heart disorders, e.g.,
CC ischaemic heart disease, cardiovascular disorders, such as ischaemic
CC heart disease. TANGO-139, 125, 110 and 175 molecules are useful to treat
CC renal (kidney) disorders, such as glomerular disease (e.g., acute and
CC chronic glomerulonephritis), TANGO-175 is useful to treat uterine
CC disorders, hyperplasia of the endometrium. TANGO-110 is useful to treat
CC spleen, e.g., the fetal spleen, associated diseases and disorder. TANGO-
CC 125 treats prostate disorders, such as inflammatory diseases, Crohn's
CC disease and tumours. TANGO-139, 125, 110, 175 or WDNM-2 are useful for
CC treating proliferative disorders, inflammatory disorders. TANGO-175, or
CC WDNM-2 activity also include apoptotic disorders, rheumatoid arthritis,
CC systemic lupus erythematosus, insulin-dependent diabetes mellitus,
CC immune-related disorders, e.g., immunodeficiency disorders, viral
CC disorders, cell growth disorders, e.g., cancers and inflammatory
CC disorders and apoptotic disorders. The nucleic acids of the invention
CC are used in gene therapy. The present sequence is human sperm-coating
CC protein (SCP)-like domain consensus sequence related to human TANGO
CC protein.
XX
SQ Sequence 162 AA;

Query Match 21.1%; Score 303; DB 23; Length 162;
Best Local Similarity 39.1%; Pred. No. 2e-20;
Matches 63; Conservative 25; Mismatches 53; Indels 20; Gaps 5;

```
Qy      52 QREIVNKHNELRRAV-----------SPPARNMLKMEWNKEAAANAQKWANQCNYRHSN 99
           | ||:|||| |: |             ||| ||  | || | |  || ||||| | | :
Db       2 QDEILNKHNDFRQQVGRGLETRGNPGPQPPASNMNPMVWNDELAQIAQNWANQCIYDHHD 61

Qy     100 PKDRMTSLKCGENL--YMSSAPS--SWSQAIQSWFDEYNDFDFGVGPKTPN---AVVGHY 152
           :  : |:|:  : |:| :  :||  || |::|  |:::           | |||
Db      62 CCRNHSKYQYGQNIAWWSSTANNYWNWSSMIQMWYNEVKDYNYNWNTCQGGNNFMVCGHY 121

Qy     153 TQVVWYSSYLVGCGNAYC-PNQKVLKYYYVCQYCPAGNWAN 192
           ||:|| ::: :||| : |  |    |:|||| ||| ||: |
Db     122 TQMVWRNTFEIGCGRSICYCNNNWHKWYYVCNYCPRGNYMN 162
```

Search completed: March 14, 2003, 05:40:27
Job time : 55.1124 secs